Analysis on Influence Factors of Enterprise Innovation Based on Data Correlation Analysis

Zaifu Sun, Weiwei Jiang

Qingdao Huanghai College, Qingdao 266400, Shandong Province, China

Keywords: Quantile Regression, OSL Regression, Enterprise Innovation

Abstract: The purpose of the study is to study the relationship of the influence of external environment on the enterprise innovation. We apply the Quantile Regression method to establish an empirical research model, and use the statistical data of some companies in 2008~2010 from Shanghai Zhenjiang High-Tech Development Zone, to analyse the influence of these factors on enterprises in different innovation quantile range. The results are compared with OLS regression. According to the regression results, we give the conclusion that the tech-efficiency, R&D investment and government policy are playing key roles in the innovation. And this result will help propose some policy suggestions to improve the enterprises innovation in China.

1. Introduction

Modern Endogenous Economic Growth Theory holds the viewpoint that technological progress and knowledge accumulation are two key factors for a country's economic growth, R&D and innovation are the most important factors to promote technological progress, and also important means to realize knowledge accumulation. China, is currently one of the fastest growing country in the pace of innovation. As it shown in the World Economic Forum, World patent filings grew at an annual growth rate of 5% in the past 10 years, contributed largely by China of 23%, which is shown in the figure 1.

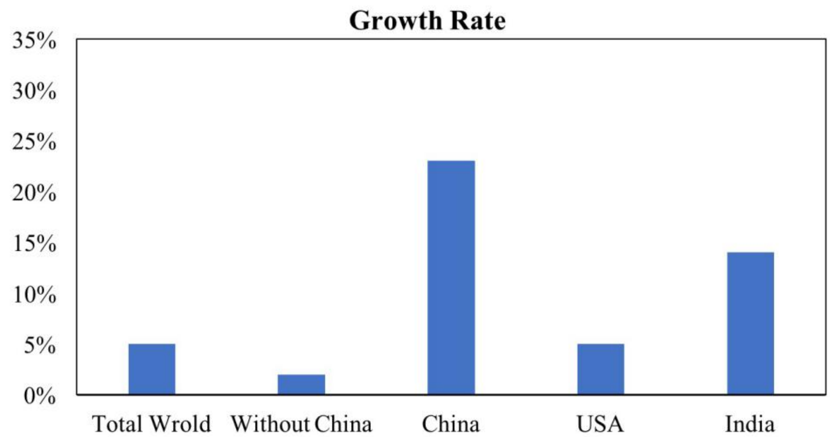

Figure 1. Global Patent Filing Growth Comparison

High-tech enterprise is the main body of research and innovation activities, and also the core force of promoting the transformation of economic growth mode and developing new industries. By improving the quality of High-tech enterprises and optimizing the relevant policies, we can solve the problem of the low efficiency of science and technology investment in China. However, High-tech Enterprises are involved in many industrial fields, with different scale and ownership form. In Shanghai, for example, the 3,111 High-tech Enterprises identified in the past three years, covering High-tech 8 Industrial technology field, employees from 11 to 44119 people, ownership including state-owned, foreign capital, private, Hong Kong, Macao joint ventures and other forms. In fact, different types of enterprises have different characteristics of innovative behavior, the

  DOI: 10.25236/ecomhs.2019.204

impact of their innovative performance factors are also significant differences, in order to improve the relevant policies in guiding the enterprise Technology Innovation behavior of pertinence and effectiveness, it is necessary to pass statistical data and reasonable mathematical tools, The factors influencing the innovation performance of different types of High-tech enterprises are analyzed. Most of the previous research literatures used OLS analysis method, because this method only considers the influence of the change of the mean value of each factor to the performance, and cannot meet the actual needs of the research, this paper studies the influence of various factors on the different innovation performance by the method of the OLS regression, and compares it with the results of the analysis.

2. Influence Factors on Enterprise Innovation

2.1. Indicator for innovation performance

Academic research on enterprise innovation performance has lasted a long time. From the narrow point of view, the enterprise innovation performance refers to the real introduction of the market of inventions, such as patents, new products, new technological process.

As far as the broad sense is concerned, the enterprise innovation performance refers to the process from producing creativity to new product creation. Success. On the measurement of innovation performance, Liu and Trevor (Liu, 2007) advocated the use of per capita new product sales revenue indicators, and to measure China's High-tech industry innovation level. Hagedoorn and Cloodt (Hageddon, 2003) argue that innovation performance is measured by R&D inputs, patents, number of citations and number of new products. However, in fact, patent indicators have obvious errors and defects. For example, many small and medium-sized enterprises because of expensive, complex process and other reasons not to apply for a patent; At the same time, the patent from the application to the authorization to occupy more time, from the patent to the actual product to invest a lot of time and energy, it is difficult to reflect the timely innovation performance. In a comprehensive sense, new product sales revenue is a more reasonable indicator of enterprise innovation performance. First of all, the new product is the enterprise research and development and innovation of the comprehensive results, more comprehensive and effective than the patent; Second, sales revenue shows that it has gained market value, and also directly reflects the competitiveness of High-tech enterprises in research and development and innovative new products. Therefore, this article uses the new product sale income to represent the enterprise innovation level.

2.2. Influence factors

The influence factors of enterprise innovation has been widely researched. A typical innovation model especially for high-tech manufacture company was proposed by Sopheon's Accolade as figure 3, which illustrated a matrix including strategy, execution for short or long term innovation.

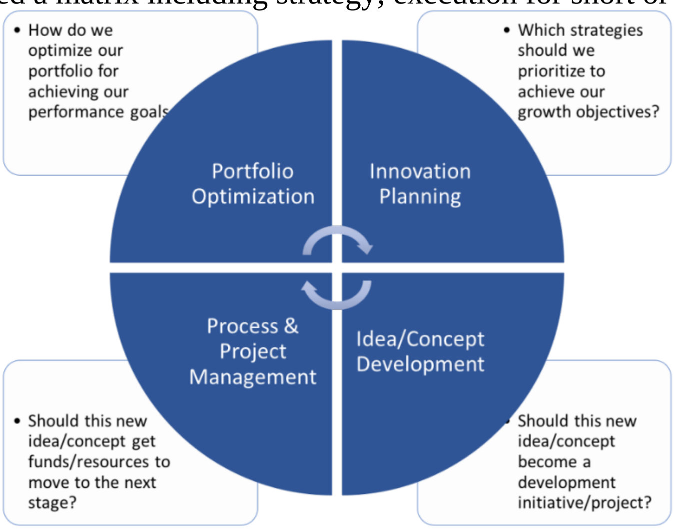

Figure 2. High-Tech Enterprise Innovation Model

Taken the Chinese national condition into consideration, we will mainly focus on the following 5 points:

1) Technology efficiency: Technical efficiency can measure the transformation relationship between innovation input and output, and it is a comprehensive reflection of enterprise production technology and management quality. This paper uses the Stochastic Frontier analysis method which is popular in recent years to calculate the technical efficiency of each enterprise, and as a regression variable into the equation, test the effect on the innovation output.

2) Enterprise scale: There are many research results on the relationship between enterprise scale and innovation behavior and performance. As early as 1942, Schumpeter through the study on the conclusion: “Market concentration and production efficiency has a close relationship between” and found that the market concentration of high, large-scale enterprise innovation effect better. Tsai (Tsai,2005) and other analysis of Taiwan Province manufacturing industry shows that there is a U-shaped relationship between enterprise scale and research and development output. This article uses the number of employees to represent the size of the enterprise, put into the equation test.

3) Government policy: From the history of science and technology development in developed countries, almost all the major achievements in the field of science and technology have traces of government subsidy. American aircraft manufacturing, Japan's large-scale integrated circuit and other industries in the process of development, the government has directly and indirectly invested in the enterprise R&D. Hu (Hu, 2001) research shows that the R&D behavior of enterprises to provide incentives to effectively promote the enterprise's research and development enthusiasm. China's central and local governments have policies to support the development of High-tech enterprises, including policies on the provision of tax breaks and subsidies for new product development. We put the “preferential policy” provided by the government into the equation and compare the output with the internal research and development input of the enterprise.

4) Ownership: Ownership Different ownership corresponds to different principal-agent relationship. The difference between the corporate governance structure and the operation mode of the enterprises with different property rights will be reflected in the output efficiency, and the difference has been confirmed in many empirical studies. Hu, through the study of the High-tech Enterprises in Haidian District of Beijing, thinks that the efficiency of non-state-owned enterprises is higher than that of state-owned enterprises in the R&D of transforming them into products.

5) Other factors: In addition to the above variables, in view of the research and development of human capital investment (with the ratio of researchers to the total number) of internal research and development investment, the ratio of net asset liability and enterprise Research and Development Organization (with the Status of scientific research cooperation), and so on has a direct impact on the innovation performance. In this paper, the research scope is limited to Shanghai Zhangjiang Hi-Tech Industrial Development Zone, which is based on the data of Shanghai Hi-Tech Enterprise panel in 2008, 2009 and 2010, so as to eliminate the interference factors such as regional difference policy difference, and the analysis conclusion is more convincing.

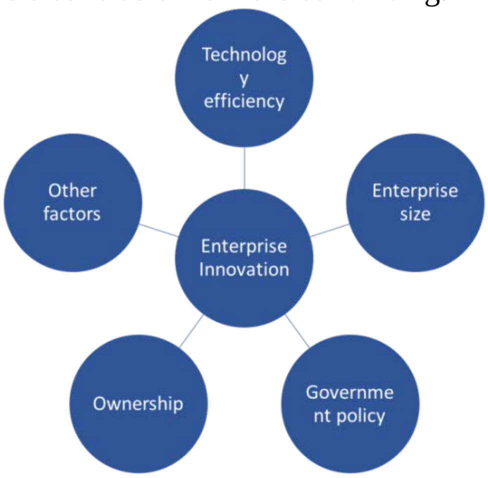

Figure 3. Innovation Impact Factors Model

3. Model

3.1. Quantile Regression

General OLS regression measures the impact of independent variable on “general enterprise”. However, this simple focus on “general enterprise” may lead to some ignorance. As Mosterller and Tukey (Mosterller, 1977), OLS calculates simply the marginal change of the dependent variable caused by marginal change of the independent variable. A more comprehensive approach is to calibrate the regression curve with respect to each quantile. Thereby, the quantile regression helps clarify the relationship between of impact factors and innovation level more comprehensively.

In comparison, the three characteristics of the regression model are as follows: First of all, OLS is not robust to the deviation of non-normal distribution, while the quantile regression is not sensitive to the anomaly value. In fact, the estimated coefficients of regression $\hat{\beta}_\theta$ will not be changed by the outliers. Secondly, the normal OLS regression only focuses on the mean change, and the distribution of the whole condition of the variable can be described by the regression of the number of digits. Finally, in contrast to the OLS regression, the regression of the number of digits is assumed to be the same distribution without residuals in each condition, so that the coefficients (elasticity) of the conditional distribution at different points in the dependent variable (innovation performance) can be estimated.

Quantile regression is put forward by Koenker and Bassett, as the following:

$$y_{it} = x'_{it} b_q + u_{qit}$$
$$Quant_\theta(y_{it} \mid x_{it}) = x'_{it} \beta_\theta \quad (1)$$

where: y_{it} denotes the innovation level, x is the impact factor, β is the parameters to be estimated, u denotes the residual. $\mathrm{Quant}_\theta(y_{it}|x_{it})$ means the θ-quantile as a function of y_{it} given x_{it}. The quantile regression parameter is obtained by solving:

$$\min_\beta \frac{1}{n}\left\{ \sum_{i,t:y_{it}\geq x'_{it}\beta} \theta\left|y_{it} - x'_{it}\beta\right| + \sum_{i,t:y_{it}<x'_{it}\beta} (1-\theta)\left|y_{it} - x'_{it}\beta\right| \right\} \quad (2)$$

3.2. Variables illustration

Taking new product revenue as the indicator, tech-efficiency, enterprise R&D investment and so on as the factor, we build the regression formula as the following:

$$IP_{it} = \beta_0 + \beta_1 Te_{it} + \beta_2 Size_{it} + \beta_3 Policy_{it} + \beta_4 HC_{it} + \beta_5 RD_{it} + \beta_6 NAD_{it}$$
$$+ \beta_7 Colla_{it} + \beta_8 OWN_2 + \beta_9 OWN_3 + \beta_{10} OWN_4 + \varepsilon_{it} \quad (3)$$

Where IP_{it} denotes the innovation level of company i in the year t. Te denotes the tech-efficiency, Size means the enterprise scale, Policy means policy tax cuts and subsidies, HC measures the ratio of the R&D personnel cost, RD means the R&D investment, NAD denotes ratio of net assets to liabilities, Colla denotes the cooperation between enterprises and external scientific research institutions condition, and Own_2 Own_3 Own_4 respectively mean foreign-funded enterprises, Hong Kong/Macao/Taiwan enterprise and domestic non-public enterprises. The contrast group is public enterprise. In addition to technical efficiency, research, development organization and ownership variables, the remaining variables are taken logarithm. The detailed explanation of these variables is attached in the following appendix. The descriptive statistics is in the Table 1.

Considering the influence of macroeconomic situation, some scholars add the statistical year as control variable to the econometric model. Because the data used in this paper is three years of statistical data, the time span is not large, and the number of virtual variables added after the model has not significantly improved, here is no longer add the statistical year to the equation in this paper statistical regression data for the Shanghai High-tech Enterprises 2008-2010 Statistics, Since every year there are enterprises in and out, excluding the sample with variable missing value, the final selection of 1840 of these samples to participate in statistical analysis.

Table 1. Descriptive statistics of all variables

Variable	Sample size	Average	Standard error	Min.	Max.
IP	1840	12.3452	3.0921	2.1309	17.8902
Te	1840	0.6789	0.1923	0.0221	0.8902
Size	1840	6.1422	1.2551	1.2233	10.3402
Policy	1840	7.2899	1.9992	2.3545	19.0898
HC	1840	-1.7652	0.7882	-4.5332	-0.1882
RD	1840	8.9921	1.6782	2.1421	18.2420
NAD	1840	1.8071	17.2234	0.0921	508.9023
Colla	1840	0.3401	0.3908	0	1
Public	1840	0.2320	0.5509	0	1
Own_2	1840	0.1580	0.4542	0	1
Own_3	1840	0.1602	0.3668	0	1
Own_4	1840	0.4208	0.5528	0	1

4. Regression Result Analysis

4.1. Regression result

The regression results are as shown in Table 2, and the leftmost one is listed as the result of random effect regression of ordinary least squares, and the right 5 column is the result of quantile regression. The results of OLS are mainly used for comparison and reference.

Table 2. Regression results

Variable	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	0.1	0.3	0.5	0.7	0.9
	IP					
Te	**6.061*****	**4.890*****	**6.543*****	**7.122*****	**6.881*****	**4.990*****
	(13.24)	(6.92)	(8.43)	(13.92)	(28.84)	(9.35)
Size	**0.753*****	**0.692*****	**0.717*****	**0.623*****	**0.702*****	**0.589*****
	(7.20)	(3.20)	(5.78)	(13.68)	(19.40)	(9.03)
Policy	**0.0898****	0.0654	0.0176	0.0356	**0.0765****	**0.0436***
	(2.48)	(0.77)	(0.33)	(1.41)	(3.17)	(2.06_
HC	**0.457*****	**0.578*****	**0.581*****	**0.439*****	**0.468*****	**0.496*****
	(8.19)	(3.12)	(8.162)	(10.03)	(13.87)	(9.55)
RD	**0.184***	0.231	0.135	**0.185***	**0.169***	**0.139*****
	(2.23)	(0.87)	(1.45)	(2.14)	(2.45)	(3.92)
NAD	**0.0885***	0.191	0.0553	0.0243	0.0058	0.017
	(2.04)	(1.23)	(1.94)	(0.72)	(0.27)	(0.71)
Colla	0.0058	**0.354***	1.39	0.0321	-0.0665	**0.078***
	(0.06)	(1.82)	(0.109)	0.38	-1.04	1.66
Own_2	-0.357	-0.4	-1.6	-0.0271	-0.008	0.0656
	(-2.57)	(-1.56)	(-0.156)	(-2.14)	(-0.19)	(0.82)
Own_3	**-0.035****	0.032	-1.23	**-0.23***	0.07	0.0364
	(-0.35)	(0.14)	(-1.643)	(-0.01)	(1.032)	(0.43)
Own_4	-0.0431	0.0408	-0.1528	-0.0018	0.0762	0.0343
	(-2.34)	(-1.45)	(-1.33)	(-0.02)	(1.22)	(0.44)
_cons	-0.016	**-0.249*****	**-1.632****	-0.475	0.155	**1.287*****
	(-0.23)	(-3.99)	(-3.01)	(-1.78)	(0.88)	(4.98)
Sample#	1840	1840	1840	1840	1840	1840
Pseudo-R^2	0.5856	0.3457	0.3567	0.4025	0.5372	0.5919

Note: *: p<0.10; **: p<0.01; ***: p<0.001.

4.2. Result analysis

According to the regression results listed in Table 2 and figure 1, we make the following analysis:

First, the impact of technical efficiency and enterprise scale are both significant and elastic in each conditional quantile point. Figure 1, the regression results show that the technical efficiency of innovation output function in every points locus on huge (due to technical efficiency in the study of the basic accord with normal distribution, this article does not do the exponential with technical efficiency and its regression coefficient is big, we only care about in the size of the different quantiles and plus or minus, so does not affect the analysis of the results). At the 50% quantile point, the improvement of technical efficiency has the greatest effect on the innovation performance. For enterprises that have achieved a certain scale of high marks and innovative output, the improvement of technical efficiency is not as significant as that of intermediate enterprises. It can be seen that improving technical efficiency is of great significance to improving enterprise innovation performance. However, the technical efficiency coefficient obtained by OLS is small in most conditional quantiles, indicating that its effect on technical efficiency is insufficient to improve the performance of innovation.

Second, compared to enterprise scale (employees) and internal R&D can see, the former on innovation performance of elasticity is larger than the latter, on the one hand, shows that Shanghai high and new technology industry to a certain extent, out of capital intensive investment growth period; On the other hand, high-tech enterprises should not lose the opportunity to expand production scale (employee number) to better improve innovation output. Internal R&D on the conditional distribution low number of coefficient is higher than high figures, such as 0.1-0.3 quantile range show in the innovation output is not high in the enterprise, increase the same amount of R&D of new products output than "star enterprise". Obviously, guiding and promoting these enterprises to increase the internal R&D investment is of greater significance to improve the overall level of innovation.

Figure 4. R&D Investment in China in 1991-2004

Third, from the perspective of the policy support of government, according to the OLS regression results preferential policies for the promotion of "general enterprise" has certain effect, but the results of quantile regression show that only in 0.7 and 0.9 two high quantiles significantly, the other all conditional quantiles are not significant. Obviously, OLS regression results have the possibility to exaggerate the promotion of corporate innovation performance by government preferential policies. The government policy is only in the promotion of enterprise performance in the innovation performance condition of the high marks.

Fourth, OLS regression results show that the R&D human capital effect on innovation performance has a certain degree of explanation, but the quantile regression shows in 0.5, 0.7, 0.9 conditional quantile level significantly, suggesting that "star" enterprises with good results for the innovation performance, increase R&D staff can obviously increase the new product sales income proportion. This indicates that, in the early stage of enterprise development, it may not be able to effectively utilize the research and development resources it has mastered because of management problems and product development. Only by developing to a certain extent and crossing the threshold of management, increasing R&D investment will bring good results.

Fifth, net asset-liability ratio under the condition of OLS quantile regression only significant and positive, the other quantile regression is not significant, but also is positive, that the improvement of enterprise net asset-liability ratio, to a certain extent, is beneficial to improve innovation level of output.

Sixth, according to the regression results of OLS, the coefficient of variable coefficient of R&D organization is greater than 0, indicating that the open enterprise can improve the innovation performance more than the closed enterprise; At the same time, the quantile regression curve is similar to the concave curve, and the cooperation between enterprises and external scientific research institutions is not conducive to the creation of innovative output at all points. For example, in a 0.1-0.3 quantile points locus on coefficient is significantly higher than other quantiles are high, this shows that the enterprise is at the early stage of market development, establish cooperation with external research institutions, colleges and universities is of great significance to promote the new product development; In quantile 0.7 appear even negative, shows innovation produces high of the enterprises and institutions of cooperation is more likely to carry out preliminary study, it inhibits the current new product sales income increase.

5. Conclusion

Using quantile regression method, this article explores the factors such as enterprise scale, the technical efficiency with the change of the innovation output of different sizes, overcome the OLS cannot reflect the differences in different innovation output conditional quantiles limitations. Through the above research, the following conclusions are drawn:

First, technical efficiency and internal R&D investment have a general and significant impact on the innovation output of enterprises. The technical efficiency is the highest in the innovation output of 0.5 point, and the internal R&D input is the most elastic at 0.1 and 0.3 of the innovation output. Net asset-liability ratio has an impact on improving innovation output, but not in general. Therefore, it is essential to encourage enterprises to strengthen management, improve labor productivity and increase internal R&D investment.

Second, firm size has a common significant influence on innovation output, and research and development cooperation at a low site flexibility than a high score in the site, which means big enterprise innovation performance is generally better, also suggests that innovation performance low sites for enterprise R&D cooperation is more important. This research conclusion is consistent with the existing theoretical research, and also accords with the innovation practice of high-tech enterprises in China. Theoretical research shows that large enterprises have advantages in systematic integration and innovation, and small enterprises have advantages in pioneering original innovation. Under the restriction of development stage, China's high-tech enterprises are mainly engaged in imitation and integrated innovation, and the advantages of large enterprises can be better played. Small businesses lack a systematic advantage and need to make up for their shortcomings through research and development cooperation. Therefore, efforts to promote the cooperation between enterprises at low points of innovation performance, especially the cooperation between innovative small enterprises and other research and development institutions, have more realistic significance for improving the performance of innovation.

Third, government preferential policies have different effect on the innovation performance of different types of enterprises. The promotion effect of government policy on innovation performance was higher at 0.1 point, with the decrease of 0.3, 0.5 points, and 0.7 and 0.9 points. It shows that the policy of tax exemption and subsidy of new and high technology enterprises has obvious effect on entrepreneurship and the newly created small enterprises. When the enterprise develops to a certain scale, it starts to innovate with its own ability, but because of the limitation of output scale and so on, it is difficult to obtain a large amount of tax credit, and the policy effect starts to weaken. This indicates that there is a bottleneck period in the development process. When the output reaches a larger scale, the tax preference will be increased and the policy effect will be strengthened gradually. Therefore, our policy support for high-tech enterprises. The paragraph should be more diversified, especially for the long-term high-tech enterprises, should not be the main policy means by tax preference.

Fourthly, the regression result of different ownership shows that the performance of foreign-funded enterprises is significantly higher than that of state-owned enterprises. On the whole, Hong Kong, Macao and Taiwan enterprises show signs of lower than the public ownership enterprises, and the non-public enterprises and the public enterprises have little difference in innovation output. This shows that ownership is not an important variable affecting enterprise innovation performance.

References

[1] Cai, W. G., & Zhou, X. L. (2014). On the drivers of eco-innovation: empirical evidence from China. Journal of Cleaner Production, 79, 239-248.

[2] Chun, D., Chung, Y., Woo, C., Seo, H., & Ko, H. (2015). Labor union effects on innovation and commercialization productivity: an integrated propensity score matching and two-stage data envelopment analysis. Sustainability, 7(5), 5120-5138.

[3] Dang, J., & Motohashi, K. (2015). Patent statistics: A good indicator for innovation in China? Patent subsidy program impacts on patent quality. China Economic Review, 35, 137-155.

[4] Dong, Y., Wang, X., Jin, J., Qiao, Y., & Shi, L. (2014). Effects of eco-innovation typology on its performance: Empirical evidence from Chinese enterprises. Journal of Engineering and Technology Management, 34, 78-98.

[5] Guan, J., & Yam, R. C. (2015). Effects of government financial incentives on firms’ innovation performance in China: Evidences from Beijing in the 1990s. Research Policy, 44(1), 273-282.

[6] Hagedoorn, J., & Cloodt, M. (2004). Measuring innovative performance: is there an advantage in using multiple indicators? Research Policy,32(8), 1365-1379.

[7] Hu, A. G. (2001). Ownership, government R&D, private R&D, and productivity in Chinese industry. Journal of Comparative Economics,29(1), 136-157.

[8] Mosteller, F., & Tukey, J. W. (1977). Data analysis and regression: a second course in statistics. Addison-Wesley Series in Behavioral Science: Quantitative Methods, 30(December), 422-424.

[9] Tsai, K. H., & Wang, J. C. (2005). Does R&D performance decline with firm size? A re-examination in terms of elasticity. Research Policy, 34(6), 966-976.

[10] Woo, C., Chung, Y., Chun, D., Han, S., & Lee, D. (2014). Impact of green innovation on labor productivity and its determinants: An analysis of the Korean manufacturing industry. Business Strategy and the Environment, 23(8), 567-576.

APPENDIX

1. Definition and measure methods of variables

Variable	Notation	Unit	Definition
Innovation performance	IP	Exponential	New product revenue (take logs)
Technology efficiency	Te	NA	Obtained by stochastic frontier analysis
Enterprise size	Size	Exponential	Total employee minus R&D (take logs)
Preferential policy	Policy	Exponential	Tax cut + subsidy (exponential)
R&D human capital	HC	Exponential	R&D employee/ total (take logs)
R&D investment	RD	Exponential	R&D investment (take logs)
debt to net worth ratio	NAD	Exponential	End of year liabilities/ (year-end assets-year-end liabilities) (take logs)
External collaboration	Colla	Virtual variable	Yes: 1; No: 0
Foreign enterprise	Own_2	Virtual variable	Yes: 1; No: 0
HK enterprise	Own_3	Virtual variable	Yes: 1; No: 0
Domestic enterprise	Own_4	Virtual variable	Yes: 1; No: 0